Language Learning & Technology January 2002, Vol. 6, Num. 1
http://llt.msu.edu/vol6num1/deridder/ pp. 123-146

VISIBLE OR INVISIBLE LINKS: DOES THE HIGHLIGHTING OF HYPERLINKS AFFECT INCIDENTAL VOCABULARY LEARNING, TEXT COMPREHENSION, AND THE READING PROCESS?

Isabelle De Ridder
University of Antwerp

ABSTRACT

This article investigates how the signaling-mode of electronic glosses in online texts (i.e., presented digitally on a computer screen) influences the user's reading process, incidental vocabulary learning, and text comprehension. Indeed, does the fact that hyperlinks with dictionary definitions are visible (i.e., highlighted) or invisible affect the foreign language learner's look-up behaviour and as a consequence the possible learning outcome? Furthermore, the article addresses the question whether the type of reading task (general vs. specific) affects the learner's use of these links and the amount and quality of the language learned. The article discusses empirical research conducted in an attempt to address these questions. The results indicate that when reading a text with highlighted hyperlinks, readers are significantly more willing to consult the gloss. However, this increased clicking does not slow down the reading process, does not affect text comprehension, and does not increase the vocabulary learned incidentally. The reading task does not seem to alter the clicking behaviour of the students but seems to influence the reader's vocabulary learning: A content-oriented reading task decreases the reader's attention for vocabulary.

INTRODUCTION

Current technologies in language learning allow student-users to consult translations, dictionary definitions, grammatical explanations, and cultural information at the simple click of a mouse. The availability of this kind of additional, often multimediatic, information is considered to be one of the pre-eminent advantages of language learning via computers, and consulting any of these extras is no longer seen as a major interruption of the language-learning activity. Research has concentrated on the effectiveness of this supplementary information and has evaluated whether these annotations improve, for instance, text comprehension or actual language learning (e.g., vocabulary learning). In general, they were found to be beneficial to several aspects of language learning (e.g., Brett, 1997, 1998; Chun & Plass, 1996, 1997; Hulstijn, Hollander, & Greidanus, 1996; Lomicka, 1998). However, although many things have been said about *what* should appear on the screen to obtain better results in language learning, *how* these features should appear to the learner-user is still under investigation. The *how* question is often considered to be a simple design question, independent of the learning process. But is this truly the case? Is the layout of the screen an autonomous issue, separate from the learning that is going on? In reading research, it has been suggested that the presentation-mode of a text on paper matters, and affects the cognitive aspects of text processing: variables such as comprehension and reading speed are said to be influenced by typefaces, margins, line length, font size (Frenckner, 1990, cited in Muter, 1996). It is still unknown to what extent findings from paper media can be extended to electronic media (Muter, 1996), but it has become clear through empirical research (see Dillon, 1992, for a review) that features such as the use of colour, screen size, interline spacing, and size of characters can play a role in the optimisation of reading from the screen and thus the learning that accompanies this reading.

Within the entire spectrum of all possible on-screen features that could have an impact on the amount and the quality of language learning, I chose to focus on the signaling-mode of glosses. Indeed, the primary aim of this study is to evaluate whether the way in which the software indicates that glosses are available influences the learners' willingness to consult the gloss. Moreover, this research concentrates on how this affects the readers' language learning (e.g., vocabulary learning) and overall text comprehension. In software designed to improve foreign or second-language readings skills, one might opt for highlighting difficult words, thus indicating a link with the provided extra information. When activating the hyperlink, these additional annotations can then appear in a pop-up window or in a separate window at the bottom or the top of the screen. More concretely, this highlighting -- which in itself is nothing but a visualisation of the hyperlink -- could be the use of boldface type or a font colour different from the text colour, for instance, with additional underlining of the word. Another option would be not to use highlights but to insert invisible links; a choice that would result in a text in entirely the same type and font colour. To summarise, a differentiation is made between the actual gloss (which is a content-related issue, i.e., the translation and/or the dictionary definition that appears when the hyperlink is activated) and the hyperlink (which is a technical feature that enables the gloss to appear). Hyperlinks can be visible or invisible: the highlights are the means by which the hyperlink is visualised (in itself a graphical issue). See Appendix A for an example of the two discussed interfaces.

I would like to stress that the effectiveness of the (electronic) gloss itself is not at issue here, since, as mentioned above, the influence of marginal glosses on vocabulary learning and text comprehension has been studied intensively. What is questioned here is whether the highlights -- used as a signalling-mode, a graphical indication of a hyperlink leading to additional information -- have an impact on the reading and/or language learning process. However, it is also taken into consideration that the possible effects of visible (i.e., highlighted) or invisible links might be affected by the reading task (general vs. specific) involved. The study hereby presented addresses these issues by discussing empirical data collected from an experiment conducted at the University of Antwerp (Belgium).

RESEARCH AIMS AND RATIONALE FOR THIS STUDY

Reading in a second language is considered to be a meaningful language-learning activity, be it on paper or on screen. Indeed, while attempting overall text comprehension, learners interact with different types of cultural, semantic, and syntactic information that can be processed and possibly learned/remembered. Moreover, many second-language specialists see reading as a pre-eminent means of acquiring new vocabulary (e.g., Krashen, 1989). This kind of vocabulary learning, where the vocabulary is "picked-up" during normal L1 or L2 reading activities, is referred to in the research as "incidental" vocabulary learning (e.g., Hulstijn et al., 1996; Nagy, Herman, & Anderson, 1985; Nation, 1990; Sternberg, 1987). It is called "incidental" because the major purpose for the interaction with the particular environment or material is not to learn words, but to understand the message of the text and to build up a coherent text base. While reading, learners seem to "guess" (Huckin & Coady, 1999), "reconstruct" or "derive" (Sternberg, 1987; Sternberg & Powell, 1983) the meaning of unknown words from the context. Whenever the context fails to deliver the correct meaning of unknown words (when the context contains too many unknown words or whenever the reader lacks the adequate cultural, world or domain knowledge [Drum & Konopak, 1987]) the use of a dictionary or of marginal (electronic) glosses is one of the factors that can promote pick-up rates (Ellis, Tanaka, & Yamazaki, 1994; Hulstijn, 1992; Knight, 1994).[1] It is therefore of great importance that students make the utmost use of the dictionary definitions the software provides.

However, a study by Black, Wright, Black, & Norman (1992) indicated that readers use electronic definitions only for short-term purposes and that long-term retention is almost non-existent. Moreover, it has also been suggested that inferred meanings are remembered better than given meanings (Hulstijn, 1992) and that inferring word meanings from context is still one of the most effective strategies for learning new words. Advocates of this kind of vocabulary learning rely on the theory of "cognitive depth"

and high mental effort: the more actively we work out a solution to the problem, the more likely we are to store this information permanently (Craik & Lockhart, 1972). Hulstijn et al. (1996) therefore suggest that dictionary use should remain "well-determined."

Summarising, the software programme should provide the learner with dictionary definitions in order to support the reader when the reading becomes problematic, but should also prevent the reader from excessive clicking (called "click happy behavior" by Roby, 1999, p. 98) which could lead to a more superficial, short-term learning. Therefore, the signalling-device of glosses plays an important part. Indeed, on the one hand, a condition in which the hyperlinks are highlighted could incite the users to click excessively, which could lead to a more temporary vocabulary acquisition. On the other hand, a condition where the hyperlinks are invisible could encourage the students to make a more careful use of the glosses and rely more on the context to figure out the meaning of unknown words (which in turn, could lead to more thorough retention of the vocabulary).

Nevertheless, in another line of thinking, it is also possible that the highlights indicating the hyperlinks simply attract attention to the words in question, which could have a positive effect on vocabulary acquisition. Several studies seem to indicate that when a word is made salient, this has a positive influence on its acquisition. In a study on listening comprehension, Brett (1998) for instance posits that if language items are made salient in any way (e.g., by exercises), they are remembered better than others. Brett did not focus on reading and vocabulary learning and therefore did not address the question of comprehension of the whole text. In reading research, some studies also indicate that when words are made salient, they are remembered better (Chun & Plass, 1996, 1997; Liu & Reed, 1995). Does this imply that in a condition without highlights (a graphical tool to make words salient), vocabulary learning would immediately become more difficult? And what would this mean in the long-term?

Some research has looked into the effects of how the users' willingness to consult an online dictionary definition can be increased. A study by Black et al. (1992; similar but separate from the one mentioned above), involving 20 subjects, suggested that if the glosses are indicated by a small black spot, displayed as a superscript immediately after the glossed word, it increases the students' willingness to consult this gloss. However, this study did not look into which effects these increased consultations might have on the reading process, actual vocabulary learning, and text comprehension.

As far as the reading task is concerned, research has established that the way in which a text is read is strongly influenced by the readers' study goals (e.g., text summarisation versus knowledge acquisition, Schmalhofer & Glavanov, 1986). It is therefore not unlikely that a given reading task might influence the students' use of either visible or invisible links, which might be reflected in the learning and/or reading outcomes. Yet another small-scale study by Black et al. (1992) could not establish that the involvement of a particular reading task (reading for gist or for detailed knowledge acquisition) differentiates the readers' clicking behaviour. The five studies by Black et al. (1992) reported on in the single article relate to first-language technical-vocabulary acquisition.

The present study aims to clarify the following issues: When difficult words are highlighted in a software reading programme, does this incite the readers to click intensively, and, if so, does this have a positive effect on the vocabulary acquisition? Or does this lead to excessive clicking, which only results in short-term vocabulary retention? Or does the simple fact that the highlights make some words in the text salient increase their chances of being better candidates for storage in long-term memory? Other studies have looked into the signaling-mode of glosses and have addressed the question of salience, but these studies did not concentrate on the combined effects of marking a hyperlink (i.e., highlighting) and the reading task (general vs. specific) on vocabulary acquisition (short-term and long-term retention), text comprehension, and reading process (reading time, clicking behaviour). Indeed, does increased clicking slow down the reading? Or does reading and rereading a higher number of glosses take up as much time as the combined action of consulting glosses and deriving a word meaning from its context? Moreover, is

it the case that when students consult more dictionary definitions, this also leads to a better understanding of the remainder of the text? Research does tend to suggest that a strong correlation exists between vocabulary and comprehension (Freebody & Anderson, 1983a, 1983b). However, perhaps it is the case that the constant interruption of the reading process results in the construction of a less coherent text base.

A small-scale, introductory study on the subject (De Ridder, 1999; De Ridder, 2000; De Ridder & Van Waes, 2000), including 17 second-year university economics students, revealed that these particular foreign language readers clicked significantly more in a condition where the text contained highlighted hyperlinks, and were thus more inclined to look for extra information (here dictionary definitions and translations) than in a condition with invisible links. This clicking behaviour did not slow down the reading process, nor did it affect text comprehension. It did have a significantly positive effect on the amount of vocabulary incidentally learned from the text. Nevertheless, this positive effect seemed to have disappeared in a delayed vocabulary test. One of the major advantages of this first study was that it made it possible to test the experimental design and the quality of the instrumentation (more appropriate and more in-depth testing was imperative). It also led to promising results worthy of more ample investigations.

HYPOTHESES

Expected Effect of Marking and Reading Task on Clicking Behaviour

Marking Increases Clicking The experimental research of Black et al. (1992) indicates that if words in an online text (i.e., presented digitally on a computer screen) are marked with a black spot behind the word, they attract the readers' attention. If the willingness to consult a gloss increases by using a small black spot, it can be hypothesised that the blue font colour and underlining would have the same effect on the subjects involved in the present study. Moreover, the previous, introductory study suggested that readers do click significantly more in the marked (or highlighted) condition than in the unmarked one.

A Specific Reading Task Decreases Clicking Research by Schmalhofer & Glavanov (1986) indicated that a text is read differently according to the reading task. I therefore hypothesised that the reading task would also change the clicking behaviour of the students. However, other research, by Black et al. (1992) for instance, suggested that a different reading task does not imply a difference in the consultation of dictionary definitions. The specific reading task used in the present study, however, was limited to strict time-constraints. It was postulated that the orientation towards the comprehension of the text, together with a slight time pressure, would incite the students to click only when absolutely necessary (i.e., when the unknown vocabulary would hinder the comprehension of the gist of the text).

Expected Effect of Marking, Reading Task and Time on Incidental Vocabulary Learning

Marking Increases Incidental Vocabulary Learning Several studies indicate that when words are made salient, this has a positive influence on their acquisition (e.g., Brett, 1998). Moreover, the more glosses are consulted, the better this might be for vocabulary acquisition (see, e.g., Hulstijn, 1992, who established a positive influence of glosses on incidental vocabulary learning). The results of the introductory study tended to confirm this argumentation. I therefore hypothesised that in the marked condition, students would pick up significantly more words than in the unmarked condition (at least in the short-term).

A Specific Reading Task Decreases Incidental Vocabulary Learning If it is the case that the specific reading task decreases the students' willingness to click, and if it is established that increased clicking results in better vocabulary learning, then this might well cause a slight drawback on vocabulary learning.

Marking Leads to a More Superficial, Short-Term Retention of Vocabulary It is more than likely that the results of the delayed vocabulary test would be significantly lower than those immediately after reading, both in the marked and the unmarked condition. However, if, as previous research suggested

(Black et al, 1992), readers tend to use electronic glosses for short-term purposes only and if in the marked condition readers are more inclined to consult definitions, then it could be expected that the results decrease significantly more in the marked than in the unmarked condition (i.e., an expected interaction effect between time and marking).

Expected Effect of Marking and Reading Task on Text Comprehension

Marking Decreases Text Comprehension Up until now, similar investigations (e.g., Black et al., 1992) were limited to the aspect of language "acquisition." To my knowledge, no study has looked into the effects of visible links on text comprehension. I hypothesised that in the marked condition, the attention of the learners would be drawn towards the highlighted words, which would distract them from the overall text and its meaning. It is indeed possible that the constant interruption of the reading process in the marked condition (more intensive clicking) would hinder the students from building up a coherent representation of the text.

A Specific Reading Task Increases Text Comprehension Since the specific reading task was much more oriented towards comprehension, the group with this particular task was expected to perform better on text comprehension than the group with the general reading task.

Expected Additional Effects of Marking

Marking Does not Slow Down the Reading Process As far as reading time is concerned, I originally assumed that if students click more in the marked than in the unmarked condition, the reading process would slow down. However, in the introductory experiment no difference in overall reading time could be established. I expected to confirm these findings in the present experiment.

Marking has a Negative Effect on the Results of the Free Recall The students of the group with the specific reading task performed a free recall. Its results were expected to be identical to the ones from the overall comprehension test, since the free recall is considered to be a fully integrated part of it. Highlighting could thus have a negative effect on the results of the free recall, because excessive clicking could lead to a poor construction of the text base.

Marking Positively Influences the Results on the Search-and-Find Question The group with the specific reading task also completed additional search-and-find questions. As far as these questions are concerned, the students might perform better in the marked condition, since highlighted hyperlinks might help in skimming a text. For instance, if the students are asked to find the four operations a diamond undergoes when it is cut, they can hypothesise that either the main category (indicated by the word *opérations*) or the subcategory (being the operations, i.e., *le clivage, le sciage, l'ébrutage, le facettage*) might be highlighted. They can start skimming the text while concentrating on the highlighted words, which would then make it easier to find the answer to the question. Previous research could not present conclusive evidence to prove that subjects find a target option faster in a highlighted display than in a display without highlighting. Some researchers favour the highlighted condition and others the non-highlighted condition (for an overview of this line of research, see Fisher and Tan [1989]).

Marking Negatively Influences the Concentration Level of the Reader It is not implausible that the blue highlights also have an impact on the students' level of concentration. In fact, in the interviews concluding the introductory experiment, the subjects involved even suggested that the highlights distracted them. The highlighted hyperlinks either attract the readers' attention and by doing so increase the interaction with the text and thus intensify the reading, or they lead to a more superficial, "click happy" behavior, where text interaction would be minimal. If students do click excessively in the marked condition and if indeed this has a negative effect on long-term vocabulary acquisition and overall text comprehension, then the concentration level of the students in the marked condition might well be significantly lower than in the unmarked condition. Appendix B, Table 11 offers an overview of the hypotheses and expectations of the experiment.

METHOD

Subjects

Sixty second-year economics students (university level; 26 females, 34 males) voluntarily participated in this experiment. These students were between 19 and 21 years of age and none of them had participated in any previous experiments. All of them were computer literate and were native Dutch speakers. None of them were bilingual but all had a fairly advanced level of French (9 years of French as a foreign language). Moreover, they were just finishing up a course of business French incorporated in their curriculum (30 hours/academic year).

Design

I randomly assigned the subjects to two groups of 30 students each, one with a general reading task, the other with a specific reading task. Both groups were subjected to a Latin-square design, as clarified in Table 1. A Latin-square design controls for text and condition order.

Table 1. Design of the Experiment

	group	reading session 1		reading session 2	
Latin-square design 1: general reading task	1 (*N*=7)	Text 1	marked	Text 2	unmarked
	2 (*N*=8)	Text 2	marked	Text 1	unmarked
	3 (*N*=8)	Text 1	unmarked	Text 2	marked
	4 (*N*=7)	Text 2	unmarked	Text 1	marked
Latin-square design 2: specific reading task	5 (*N*=7)	Text 1	marked	Text 2	unmarked
	6 (*N*=8)	Text 2	marked	Text 1	unmarked
	7 (*N*=8)	Text 1	unmarked	Text 2	marked
	8 (*N*=7)	Text 2	unmarked	Text 1	marked

Material

Reading Materials The students involved in the experiment read two glossed French economic texts, comparable in length (about 2,000 words each), grammatical difficulty, and vocabulary load. The texts had been selected for the introductory study mentioned above, after a pilot study involving four texts, 28 students and two faculty members of French. Within this pilot study, the texts were evaluated for interest and difficulty level: the two texts with the most similar score were included in the study. The first text dealt with the diamond industry in Antwerp; the second one was about human resources in business. A word was glossed whenever one student of the pilot study failed to know the Dutch translation of this word in a vocabulary test taken after reading. This procedure lead to the creation of 109 glossed words in the first text and 116 in the second one, which is about 5-6% of the total amount of words. In the existing literature, there is no consensus on how many unknown words a text may contain in order not to disturb the global comprehension level or the learning of vocabulary: West (1941) speaks of 2%, whereas for instance Holley (1973) refers to 7%. Both studies did not include the help of glosses.

The students were asked to read the texts online. They could easily access the glosses by clicking on the defined word. A pop-up window with a Dutch translation and a French definition (separated with a horizontal line) would then appear. I chose to offer the translation and the foreign language definition, since preferences for one or the other are said to be highly individual (Jacobs, Dufon, & Fong, 1994). The pop-up window did not cover up the portion of the text in which the glossed word was found (see Roby, 1999; Stark, 1990). Context-bound explanations were given first; more general information was given at the end between brackets. This difference was explained to the students before starting their reading (see Widdowson, 1978).

Two versions of each text were created: a marked one and an unmarked one. In the marked condition the glossed words were in blue and underlined. In the unmarked condition the hyperlinks were invisible (i.e., the glossed words were typed in black and not underlined; see Appendix A). The glosses remained identical in both conditions. The texts appeared in a black Times New Roman font, 12 points, on a white screen. The screen design was based upon the "Recommendations for Basic Typography and Spatial Factors" of Scott Grabinger & Osman-Jouchoux (1996, p. 194-196). These authors give recommendations for screen design, based upon both print and computer screen research. In the investigation carried out, the following aspects were divergent from these guidelines: the full justification (instead of left justification) and the line length (83 characters). Scott Grabinger and Osman-Jouchoux propose a line length of 60 characters, while other authors (e.g., Dyson & Kipping, 1998) find an increase in reading rate with a greater number of characters per line.

Instrumentation and Observation

Clicking Behaviour and Total Reading Time An *Internet Explorer* specific java-script made it possible to register how much time the students spent on reading and on clicking. With these log files it could be determined which glosses the students consulted and for how long they consulted them. The students were shown how the glosses worked before starting their reading and they were explicitly told to close each pop-up window after consultation.[2]

Vocabulary Test After having read each text, the students took an unexpected vocabulary test of 38 items. All of these items were chosen on the basis of the original pilot study and the glossed words. For every word of the vocabulary test, the students were asked whether they thought they knew the word a) already before reading the text or b) from reading the text. In this way, the individual pre-knowledge of the students could be estimated. I deliberately chose not to include a pre-test in the experiment because in this particular kind of experiment, testing the students on vocabulary knowledge before they start reading might direct their reading towards vocabulary. Within this learning indication survey, an "I don't know" option was also included. This same technique was applied by Hulstijn et al. (1996) in their study on the combined influence of frequent occurrences and the use of dictionaries/marginal glosses on incidental vocabulary learning. One week after having read the text, the students took a delayed vocabulary test (the same items but in a scrambled order) in order to measure any long-term retention of the vocabulary eventually acquired.

The two texts also contained four non-existent words each. To these words, I attached a self-made, acceptable, context-bound meaning.[3] In the introductory experiment, only two of these words per text were included, but an item selection test (included in Cronbach's reliability test) proved that these words were very good discriminators, hence, the decision to increase their number. I deliberately avoided working exclusively with non-existent words because of the artificial nature of this procedure. The four non-existent words of the present study were also used as a means of verification. Since it was impossible for the students to know these words, these contributed to the clarification of the students' clicking behaviour and the learning indications that were included in the tests. When scoring the vocabulary test, I did not take into account the words that at least 70% of the students said they knew before reading the text (indicated in the learning indication survey) and answered correctly in both of the tests.[4]

Text Comprehension Test After having read the text, all of the students took a comprehension test of 8 multiple-choice questions and 13 open-ended questions. The reliability scores for this test were rather low, but could be improved by leaving out some of the items of the test (item selection test, *N* of cases = 60, Text 1, α = .7021, *N* of items = 11, Text 2, α = .7351, *N* of items = 13). The group with the specific reading task started with a search-and-find task, which was followed by a reading instruction that specifically described the comprehension test that was to be taken afterwards: the general comprehension test mentioned above (multiple choice and open-ended content questions) and a summarisation under the form of a free recall. The tasks in question were chosen on the basis of the work of Wesdorp (1981),

Kintsch (1998, p. 295), and Nuttall (1996), who give an overview and a critical evaluation of assessment techniques for reading comprehension. Several studies on the teaching of reading (e.g., Benítez, Castrillo, Cerezal, & Suárez, 1988; Nuttall, p.150) propose skimming through a text (stimulated by search questions) before starting the actual reading as a successful strategy. For an in-depth insight in search-and-do and other possible reading strategies, see for instance Fyfe & Mitchell (1985).

The search-and-find task of this study contained four questions where one or more items had to be found: three questions could be found literally in the texts, and for the fourth one, the students had to establish a relationship between information contained in two (not necessarily consecutive) paragraphs. The reliability scores for the search-and-find task were rather low, but could be improved by leaving out 1 item in each test (see item selection test, *N* of cases = 30, *N* of items = 9, Text 1, α = .6979, Text 2, α = .7077,).

The evaluation procedure of the free recall occurred as follows: For every separate part of the texts, main and secondary ideas were identified. Whenever a student was able to reproduce a main idea and two or three secondary ideas, depending on the level of difficulty of that particular part of the text, the student received 3 points. If s/he only mentioned the main idea without the necessary secondary ideas, the student was assigned 2 points. Only reproducing the subsidiary ideas entitled the student to 1 point. Students could obtain a maximum of 34 points for text 1 and 32 for text 2.[5]

Concentration Level of the Students To establish a possible difference in the students' concentration during the reading of the marked versus unmarked text, a small attention test for the group of the specific reading task was incorporated. During the online reading sessions, a red rectangle (4 to 3 cm, java-script) randomly appeared (with a maximum of 5 times in the 25 minutes of reading time). The students were explicitly asked to click on it as fast as possible in order to make it disappear. I expected that the longer it took the students to react, the higher their concentration level was and the higher their interaction level with the text they were reading. A similar technique was used in research on writing in order to establish the subjects' attention capacity (Kahneman, 1973; Kellog & Mueller, 1993).[6] The goal of this test is to disturb the students in their reading, and the longer it takes them to withdraw their attention from the text, the more they are assumed to be concentrating on their reading.

Additional Information Gathering All 60 students were allowed to take notes while reading so as not to disturb their normal reading process. They were not, however, allowed to use these notes during any of the tests. The process data were collected by recording all sessions with Hypercam,™ a software programme that registers every on-screen movement and files them in AVI-videofiles. These files were used as a means of verification. For all 60 students a small interview and some general questions on the student's reading habits concluded each session.[7]

Procedure

All students participated in three individual sessions. Each session was concluded with a small interview and some general questions. Before reading, the students received a general technical explanation (working of pop-ups, etc.).

The General Reading Task-Group (Latin-square 1) In the first session, the students of this particular group read one of the two texts in one of the two conditions, took a test on text comprehension and completed an unexpected vocabulary test. In the second session, they started with the delayed vocabulary test of the text of the first session, read the other text in the other condition, took a test on text comprehension and a finished with a vocabulary test. The delayed vocabulary test of the second session took place in a brief third session. The 30 students of this group received the following (general) reading task before reading:

> Read the text that will appear on the screen very thoroughly and try to understand as much as you can. After having read the text, you will receive a test on text comprehension. You can read as long as you think is necessary.

Thus, no specific time limits were set and no particular information about the testing that would follow the reading was included. Table 2 gives an overview of the development of the experiment within this (general reading task) group.

Table 2. Development of the Experiment Within the Group with the General Reading Task

activity	session 1	session 2	session 3
reading of one of the two texts in one of the two conditions (see Latin square design)	X	X	
test on text comprehension	X	X	
vocabulary test	X	X	
interview and general questions	X	X	
delayed test on the vocabulary of the text read in previous session		X	X

The Specific Reading Task Group (Latin-square 2) In the first session, the students of the specific reading task group started with a search-and-find task of one of the two texts, in one of the two conditions. Afterwards, they read the text in question, performed a free recall and took an unannounced vocabulary test and an announced comprehension test. In the second session, they took a delayed vocabulary test of the first text and read and completed the tasks of the second text. The delayed vocabulary test of the second session took place in a brief, third session.

For the search-and-find task, the students received a total of 6 minutes. Afterwards, they were assigned the following reading task, specifically mentioning the form in which they would be tested:

> You will receive 25 minutes to read the text that will appear on the screen. Afterwards, you will be asked to give an oral overview of what you have read in the text. You will also receive a comprehension test with multiple-choice and open-ended questions.

As the reader can see, this particular experimental group was subjected to strict time limits, which were pre-tested with advanced learners and were adjusted to this specific group of learners. The students had the same comprehension test as the first group and they took the same vocabulary test after the free recall.[8] In Table 3 one can find an overview of the development of the experiment within the group with the specific reading task. Activities marked with an asterisk are identical to the ones of Table 2 of the group with the general reading task.

Table 3. Proceedings of the Experiment for the Group with the Specific Reading Task

activity	session 1	session 2	session 3
search assignment for one of the texts in one of the two conditions (see Latin square design; 6 minutes)	X	X	
reading of one of the texts in one of the two conditions (see Latin-square; 25 minutes)	X	X	
free recall	X	X	
vocabulary test (*)	X	X	
comprehension test (*)	X	X	
interview and general questions (*)	X	X	
delayed test on the vocabulary of the text read in previous session (*)		X	X

RESULTS

For an overview of all established results, see Appendix B, Table 12. Within the design of this experiment, marking and time are within-subjects variables; reading task is a between-subjects variable. A within-subjects design means that each participant provides more than one response. With a between-subjects variable, every set of responses comes from a different group of subjects.

Effect of Marking and Reading Task on Clicking Behaviour

Table 4 shows the percentage of time applied to clicking. The results were analysed with marking as a within-subjects variable and reading task as a between-subjects variable.[9] The results show that marking and reading task have a significant effect on the students' clicking behaviour: marking, F (1,58)=24.292, p < .05; reading task, F (1,58)=16.331, p < .05. There is no interaction effect between the within-subjects variable and the between-subjects variable: F (1,58)=0.008, p > .05. These results show that students click significantly more in the marked than in the unmarked conditions and significantly more in the group with a general reading task than in the group with the specific reading task. Appendix C, Figure 3 illustrates these results. Marking thus increases clicking and a specific reading task decreases clicking, as was hypothesised.[10] All statistical tests were performed at the .05 level, unless otherwise indicated.

Table 4. Percentage of Time Applied to Clicking

condition	marked				unmarked				*N*
	Mean	*SD*	Min	Max	Mean	*SD*	Min	Max	
general reading task	12.15	6.50	1.11	26.82	9.44	5.05	1.85	20.16	30
specific reading task	7.70	4.00	.86	19.58	4.88	3.02	.61	11.52	30
total	9.92	5.80	.86	26.82	7.16	4.72	.61	20.16	60

GLM, repeated measures	*SS*	*MS*	*df*	*F*	*p*
effect of marking (within-variable)	228.83	228.83	1	24.290	.00*
effect of reading task (between-variable)	608.17	608.17	1	16.330	.00*
interaction effect marking*reading task	7.65	7.65	1	.008	.929

Effect of Marking, Reading Task and Time on Incidental Vocabulary Learning

Table 5 presents the results of the vocabulary tests, taken immediately after reading and then one week later. These results were analysed with marking and time (test immediately after reading and delayed test) as within-subjects variables and reading task as a between-subjects variable. The results show that on one

hand, marking has no significant effect on vocabulary learning: F (1,58)=3.70, p >.05. On the other hand, time and reading task do have a significant effect on vocabulary learning: time, F (1,58)=10.62, p <.05; reading task, F (1,58)=28.16, p <.05. There are no interaction effects: marking*reading task, F (1,58)=0.17, p >.05; marking*time, F (1,58)=2.59, p >.05; time*reading task, F (1,58)=1.30, p >.05; marking*task*time, F (1,58)=.127, p >.05). Both groups thus score significantly lower in the delayed vocabulary test. The group with the general reading task scores significantly better on the vocabulary test than the group with the specific reading task.[11] Appendix C, Figures 4 and 5 illustrate these results. Thus, marking does not specifically influence incidental vocabulary learning (contrary to what was hypothesised), whereas a specific reading task does decrease vocabulary learning (as was hypothesised). Moreover, marking does not lead to more superficial, short-term retention of vocabulary learning, since there is no significant interaction effect between time and marking (contrary to what was hypothesised).[12]

Table 5. Results on the Vocabulary Test, Immediately After Reading and Delayed Vocabulary Test

condition	marked				unmarked				
IMMEDIATELY AFTER READING	Mean	*SD*	Min	Max	Mean	*SD*	Min	Max	*N*
general reading task	70.27	18.96	31.43	97.14	66.21	13.30	36.67	88.57	30
specific reading task	53.87	15.49	17.14	80.00	48.66	16.36	25.71	80.00	30
total	62.07	19.06	17.14	97.14	57.43	17.23	25.71	88.57	60

condition	marked				unmarked				
ONE WEEK LATER	Mean	*SD*	Min	Max	Mean	*SD*	Min	Max	*N*
general reading task	66.91	15.93	33.33	97.14	65.48	11.36	28.57	85.71	30
specific reading task	47.55	16.21	16.67	73.33	46.47	15.06	23.33	76.67	30
total	57.23	18.69	16.67	97.14	55.97	16.33	23.33	85.71	60

GLM, repeated measures	*SS*	*MS*	*df*	*F*	*p*
effect of marking (within-variable)	520.91	520.91	1	3.70	.059
effect of time (within-variable)	595.39	595.39	1	10.62	.002*
effect of reading task (between-variable)	19611.48	19611.48	1	28.16	.000*
interaction effect marking*reading task	2.32	2.32	1	.017	.898
interaction effect marking*time	171.599	171.599	1	.259	.112
interaction effect time*reading task	72.81	72.81	1	1.30	.259
interaction effect marking*task*time	8.41	8.41	1	.127	.722

Effect of Marking and Reading Task on Text Comprehension

Table 6 shows the results of the overall comprehension test, that is, the multiple-choice questions and the open-ended questions. These results were analysed with marking as a within-subjects variable and reading task as a between-subjects variable. The results show that neither marking nor reading task have a

significant effect on the student's text comprehension: marking, F (1,58)=.003, p >.05; reading task: F (1,58)=.028, p >.05. There is no interaction effect between the within- and between-variable: F (1,58)=.246, p >.05. Appendix C, Figure 6 illustrates these results. Thus, marking does not decrease comprehension of the text and a specific reading task does not increase text comprehension (both contrary to what was hypothesised).

Table 6. Result of the Text Comprehension Test in Percentage

condition	marked				unmarked				
	Mean	*SD*	Min	Max	Mean	*SD*	Min	Max	*N*
general reading task	57.13	28.35	0.00	100.0	61.93	23.34	9.09	100.0	30
specific reading task	62.91	16.44	30.77	100.0	57.78	20.84	9.09	92.31	30
total	60.02	23.16	0.00	100.0	59.86	22.04	9.09	100.0	60

GLM, repeated measures	*SS*	*MS*	*df*	*F*	*p*
effect of marking (within-variable)	.799	.799	1	.003	.959
effect of reading task (between-variable)	19.968	19.968	1	.028	.869
interaction effect marking*reading task	739.54	739.54	1	2.46	.122

Additional Effects of Marking

Total Time Spent on Reading Table 7 presents the total time in seconds spent on reading, only for the group with the general reading task, since the group of the specific reading task was subjected to strict time limits. These results were analysed with marking as a within-subjects variable, showing that there is no significant effect of marking on the total time spent reading: F (1,58)=1.54, p > .05. Thus, marking does not slow down the reading process, as was hypothesised.

Table 7. Total Reading Time in Seconds in the Group of the General Reading Task

condition	Mean	*SD*	Min	Max	*N*
marked	2071.46	581.34	1113	3177	30
unmarked	1927.13	590.10	1076	3003	30

GLM, repeated measures	SS	*MS*	*df*	*F*	*p*
effect of marking (within-variable)	312481.66	312481.66	1	1.54	.224

Free Recall Table 8 shows the results of the free recall performed by the group with the specific reading task. These results were obtained by averaging the evaluations of two independent researchers who analysed the transcriptions of the oral recalls produced by the students. Agreement between the scores of both independent researchers was measured by computing the Pearson product moment correlation coefficients; for text 1 a .937** correlation was achieved and for text 2 a .974** correlation (= significant at the .01 level). For issues in measuring reliability, see Hayes and Hatch (1999). These results were analysed with marking as a within-subjects variable, revealing no significant effect of marking on the free recall: F (1,29)=1.18, p > .05. These results are thus identical to the results on the overall text comprehension test: marking does not negatively influence text comprehension.

Table 8. Results of the Free Recall in Percentage

condition	Mean	*SD*	Min	Max	*N*
marked	1.72	12.69	4.69	53.13	30
unmarked	34.11	15.09	7.81	72.06	30

GLM, repeated measures	*SS*	*MS*	*df*	*F*	*p*
effect of marking (within-variable)	85.82	85.82	1	1.18	.285

Search-and Find Question Table 9 displays the results of the search-and-find questions in percentage. These results were obtained by counting all correctly found items. The search-and-find task only applies to the students of the group with the specific reading task. I analysed these results with marking as a within-subjects variable, which shows that there is a significant difference between the results of the marked and the unmarked condition: F (1,29)=7.02, $p < .05$. Marking has a negative influence on the results of the search-and-find task (contrary to what was hypothesised), since the students score significantly better in the unmarked condition.

Table 9. Results of the Search-and-Find-Question in Percentage

condition	Mean	*SD*	Min	Max	*N*
marked	53.55	32.23	0.00	100.00	30
unmarked	74.00	25.31	0.00	100.00	30

GLM, Repeated Measures	*SS*	*MS*	*df*	*F*	*p*
effect of marking (within-variable)	6269.62	6269.62	1	7.02	.013*

Attention Test The values in Table 10 are the results of the attention test. They represent the seconds the students with the specific reading task needed to close the red rectangle on the screen. These results were analysed with marking as a within-subjects variable, which revealed that no significant difference could be established between the marked and the unmarked condition: F (1,29)=.220, $p > .05$. According to these results, marking does not negatively influence the concentration level of the students (contrary to what was hypothesised).

Table 10. Seconds Necessary to Close the Red Rectangle on the Screen (Attention Test)

Condition	Mean	*SD*	Min	Max	*N*
marked	2.54	1.15	1.25	5.75	30
unmarked	2.66	1.40	1.00	6.00	30

GLM, repeated measures	*SS*	*MS*	*df*	*F*	*p*
effect of marking (within-variable)	.217	.217	1	.220	.642

DISCUSSION AND CONCLUSION

In this section, the questions raised in the research aims and rationale section are revisited and discussed in light of the results of the empirical investigation presented above. The first question posed, which led to several other questions, was whether highlighted or visible links increase the readers' willingness to consult dictionary definitions. According to the results of the present study, the answer to this question is affirmative. The student users consulted significantly more glosses in the condition with visible links. This finding confirms previous research of Black et al. (1992) concerning first-language acquisition and shorter texts, where a black spot behind the glossed words also attracted the learners' attention and made

them click to access the provided word definition. Is this particular clicking behaviour altered by the reading task? The findings of the experiment indicate that students click significantly more in the marked than in the unmarked condition, be it in the general reading task or the specific reading task group. This might be an indication of how powerful a tool highlighting is to attract the reader's attention. However, the study also established that readers click considerably more when they are confronted with a general reading task than when they receive a specific one, while previous research (Black et al.,1992) seemed to suggest that a reading task variable does not influence the students' clicking behaviour. It is possible that the difference between the present results and those of Black et al. are due to the time limits that were imposed on the students.

Furthermore, does the fact that students spend significantly more time consulting information in the marked than in the unmarked condition influence the incidental learning of vocabulary? On the short-term-vocabulary test (taken immediately after reading) no difference was established, which means that whichever different vocabulary-learning strategy the students used in the two conditions can be considered equally effective in the short-term. It seems plausible that in the marked condition, students turn to a vocabulary-learning strategy based upon the reading of dictionary definitions, whereas in the unmarked condition they use a combination of this particular strategy and context derivation. The findings of the present experiment indicate that not using highlights thus making the glossed words less graphically salient, does not particularly influence the incidental learning of vocabulary in a negative way. In both conditions, the students seemed to have adapted their vocabulary learning strategies to the marking situation. In the delayed vocabulary test, the students scored significantly lower than in the test taken immediately after reading. However, the results indicate that in the marked condition, where students were guided more intensively by the highlights and clicked more excessively, the vocabulary loss is not greater than in the unmarked condition. The findings thus do not suggest that readers tend to use electronic glosses for short-term purposes, as the investigations of Black et al. (1992) indicated. The findings of the investigation presented also establish that the use of either strategy does not slow down the reading process, since no difference in reading time was found between the marked and the unmarked condition.

Does the reading task have an effect on the learning of vocabulary? As indicated, reading is often seen in the research as a pre-eminent means of vocabulary learning. The findings of the empirical research presented above clearly indicate that the success of this vocabulary learning depends highly on the reading task that is set. In the experiment, the specific reading task led to significantly less incidental vocabulary learning, which is more than probably due to the time pressure and the specific orientation towards text comprehension.

Did the more intense clicking of the marked condition lead to better text comprehension? Apparently not, and this might be explained by the link that seems to exist between vocabulary knowledge and reading comprehension, which was suggested by previous research (Freebody & Anderson, 1983a, 1983b). Since both conditions result in equal learning of vocabulary, both conditions also result in equal text comprehension.[13] In the experiment, text comprehension was measured in different ways: on the one hand, a general comprehension test containing multiple-choice and open-ended questions was used; on the other hand an additional free recall for the group with the specific reading task was employed. Neither of these indicated an effect of marking on text comprehension. The search-and-find question, on the other hand, was subjected to an effect of marking. These results seem to confirm part of the research on localising items in highlighted and non-highlighted displays. It is possible that the highlights confused the students and that in the unmarked condition the text structure helped them to find the items they were looking for.

Did the reading task have an effect on text comprehension? The findings of the current investigation indicate that, contrary to what was hypothesised, the introduction of the specific reading task did not lead to a more thorough understanding of the text. This might have been due to the time limits that were set.

The fact that students concentrated more on the text content when they received a specific reading task is suggested by additional information from the notes they took while reading. In these notes, they wrote down exclusively content-related items, whereas the other group (general reading task) wrote down more vocabulary.[14]

Are students more concentrated in the unmarked than in the marked condition? Is it possible for them to interact more intensively with the text in the unmarked condition? The attention test that was introduced did not reveal any differences in the concentration level of the students.

The results of the current experiment do not all confirm the findings of the previous, introductory experiment. Globally, the findings are consistent. However, the present ones are much more detailed and complete (especially on the level of text comprehension). The previous experiment did indicate a better result on the short-term vocabulary test after reading the texts in the marked condition. Nevertheless, since the internal consistency scores of the current tests were higher than the previous ones, it may be assumed that the current results are more reliable. On the other hand, the p-value of this particular effect is here .059, which is not statistically significant when a 5% level of significance is maintained. However, I believe that this value indicates that further research might not be redundant.

The fact that no significant effect could be established is perhaps due to the population of the study. Indeed, the present study was conducted with students who probably have a rather high and advanced level of French, especially when compared to other cultural backgrounds: they all had 9 years of French as a second language and live in a country where French is one of three official languages (also Dutch and German). It is possible that the results would be completely different with students at the beginning or intermediate level, coming from different cultural backgrounds. It might also be the case that the basic assumption that more intensive clicking leads to better results on the vocabulary test -- as was suggested by previous research on glosses -- is one that should be reconsidered. Indeed, when computing the Pearson product moment correlation coefficient with the results of the vocabulary test and the time applied to clicking, a .329* correlation was achieved for the marked condition and a .270* correlation for the unmarked condition (= significant at the .05 level). These are significant but very weak correlations.

Summarising, the present experiment made it possible to evaluate what influence the signalling-mode of electronic glosses has on vocabulary learning, text comprehension and the reading process. Indeed, when foreign-language learners read a text where the link with the gloss is visible (highlighted), they are more willing to consult the gloss. However, this increased clicking does not slow down the reading process, nor does it increase the vocabulary learned incidentally. On the contrary, when reading a text in a condition with invisible links, the students' clicking will be less excessive and better determined, leaving room for context derivation, which, in the long-term however, does not particularly have a positive effect on vocabulary learning. The fact of highlighting or not highlighting the hyperlink does not have an impact on text comprehension either. Apparently, the readers seem to adapt their reading strategies and vocabulary learning strategies to the screen-situation they are confronted with. The reading task then does not alter the clicking behaviour of the students since they still click considerably more when visible links are presented, even when carrying out a specific reading task. However, the reading task did influence the students' vocabulary learning: A content-oriented reading task seems to decrease the reader's attention for vocabulary.

Nevertheless, in a future follow-up investigation, some issues of the present experiment could still be improved, for instance, the way in which the study dealt with prior knowledge. I think my reasons for not including a pre-test in the experiment are defendable. The vocabulary test was based upon results with similar students and in co-operation with the students' teaching assistants. By asking the students whether they knew the word before, a means of verification was included. In the end, 30% of the words of the original vocabulary test was not taken into consideration, yet an internally consistent vocabulary test was kept with the remaining items. As far as text comprehension is concerned, some questions could not yet

be answered. For instance, no significant correlations could be found between the different text comprehension tests (i.e., the search-and-find question, the overall comprehension test, and the free recall). This would imply that somehow these tests have measured other types of comprehension. It remains to be seen exactly where the differences between the three lie. However, it should not be forgotten that the Cronbach's alpha values of the free recall were rather low and that therefore the absence of effects or significant correlations could be due to the unreliability of this particular test.

When listening to all 60 students in the interviews, I realised that different students might have experienced the different conditions in a different way. Moreover, the videofiles from the Hypercam™ reveal that within the group of participating students different reading or learning profiles might have been involved. Furthermore, the standard deviation (*SD*) of the percentage of time applied to clicking is rather high (see Table 4) which is also an indication of differences between the individual students. In a follow-up investigation, I therefore plan to include aspects of cognitive mapping in the research. The results of the current research indicate that the learners involved were rather flexible and adapted easily to the material they were confronted with. However, the question can still be raised whether the users should adapt to the design of the screen and not the other way around. Visible (highlighted) hyperlinks might be helpful for some learners while disturbing for others, as far as vocabulary learning, text comprehension and the reading process are concerned. Students whose learning style is one of external regulation and step-by-step processing might benefit more from a highlighted condition than students whose learning style is one of self-regulation and deep-processing (Vermunt & Van Rijswijk, 1987).

The kind of research that involves cognitive flexibility and cognitive profiles would undoubtedly be useful, especially in the light of today's technological developments. With the advent of Dynamic HTML for instance, the design of the screen can be instantly adapted to the learning and reading profile of the user (e.g., Godwin-Jones, 2000). The notion of "usability" (see Nielsen, 2000) is becoming a component that online educational settings can no longer ignore.

APPENDIX A

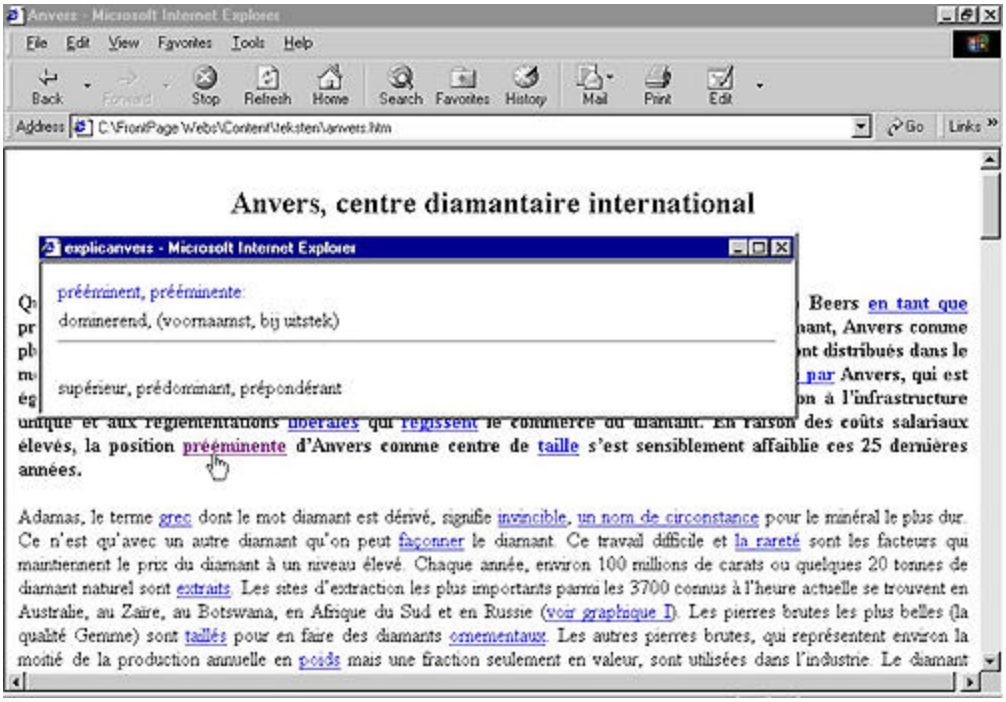

Figure 1. Screenshot of the interface using highlights (words in blue and underlined)

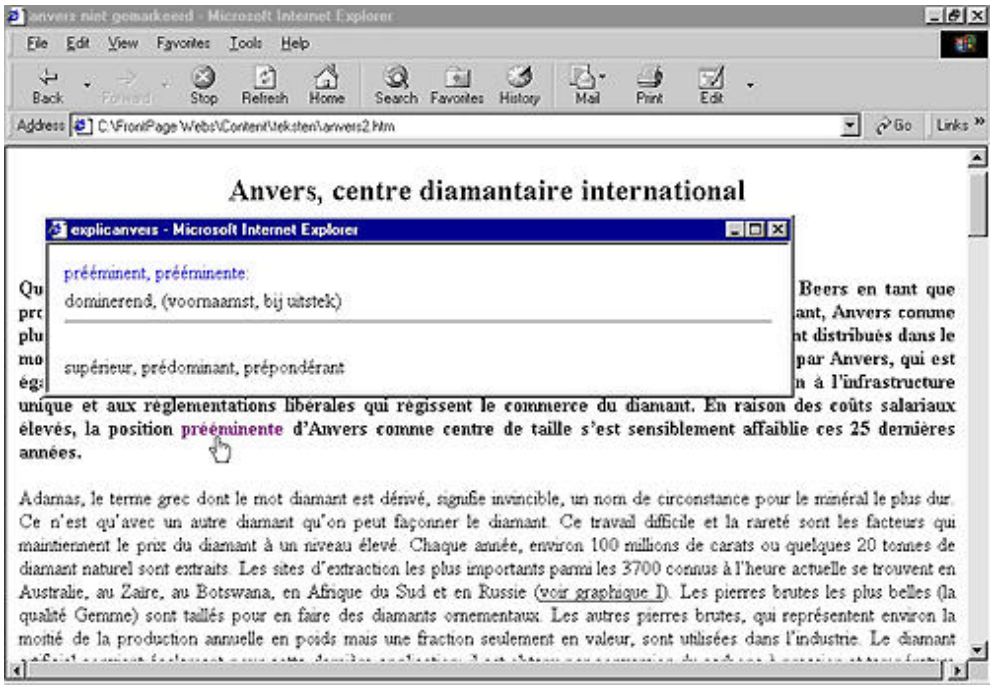

Figure 2. Screenshot of the interface without highlights (invisible links)

APPENDIX B

Table 11. Overview of the Hypothesis of Experiment 2

	effect of marking (within-subjects variable)	effect of reading task (between-subjects variable)	effect of time (within-subjects variable)
Reading process			
clicking behaviour	positive effect	negative effect	
total reading time	no effect		
Vocabulary learning	positive effect	negative effect	negative effect + interaction effect with marking
Text comprehension	negative effect	positive effect	
Additional testing			
free recall	positive effect		
search-and-find	positive effect		
attention	negative effect		

Table 12. Overview of the Established Effects of Experiment 2

	effect of marking (within-subjects variable)	effect of reading task (between-subjects variable)	effect of time (within-subjects variable)
Reading process			
clicking behaviour	positive effect	negative effect	
total reading time	no effect		
Vocabulary learning	no effect	negative effect	negative effect but no interaction effect
Text comprehension	no effect	no effect	
Additional testing			
free recall	no effect		
search-and-find	no effect		
attention	no effect		

APPENDIX C

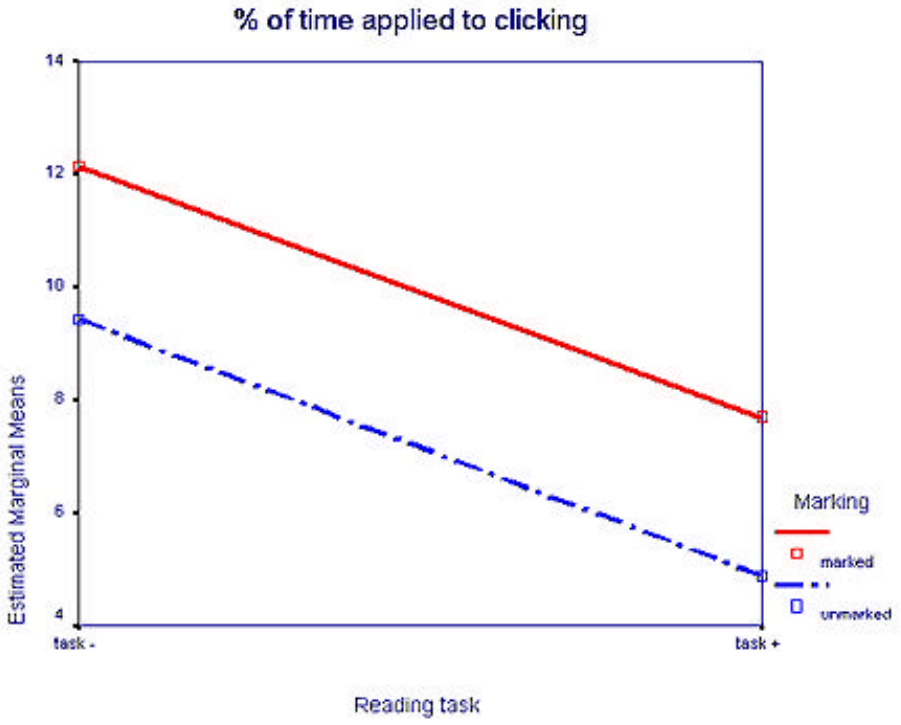

Figure 3. Percentage of Time Applied to Clicking, Within-Subjects and Between Subjects Variable

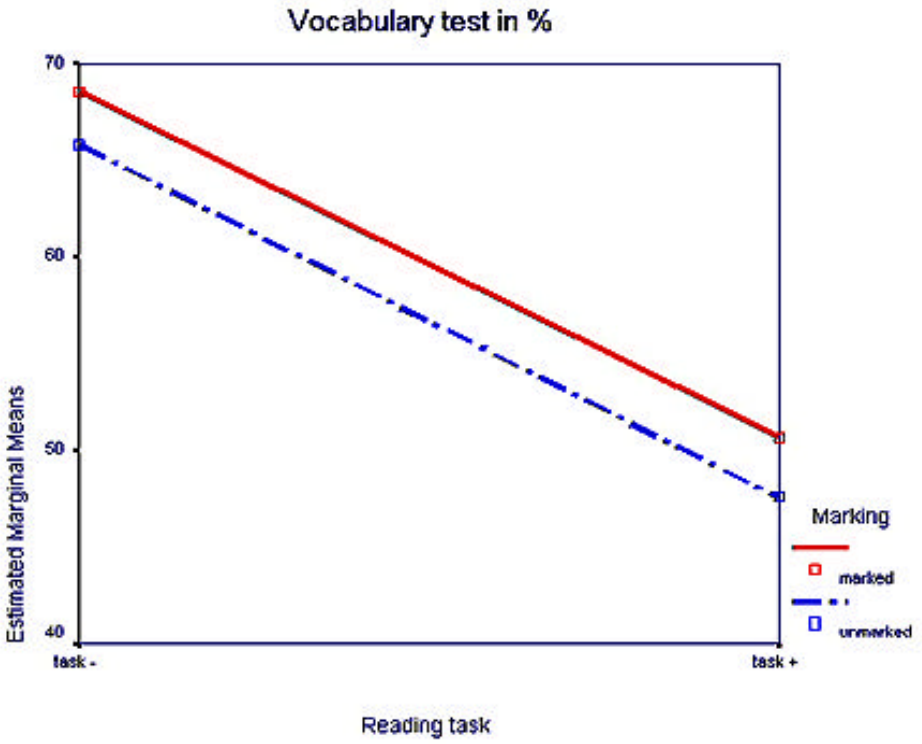

Figure 4. Vocabulary Test in Percentage, Within-Subjects and Between-Subjects Variable

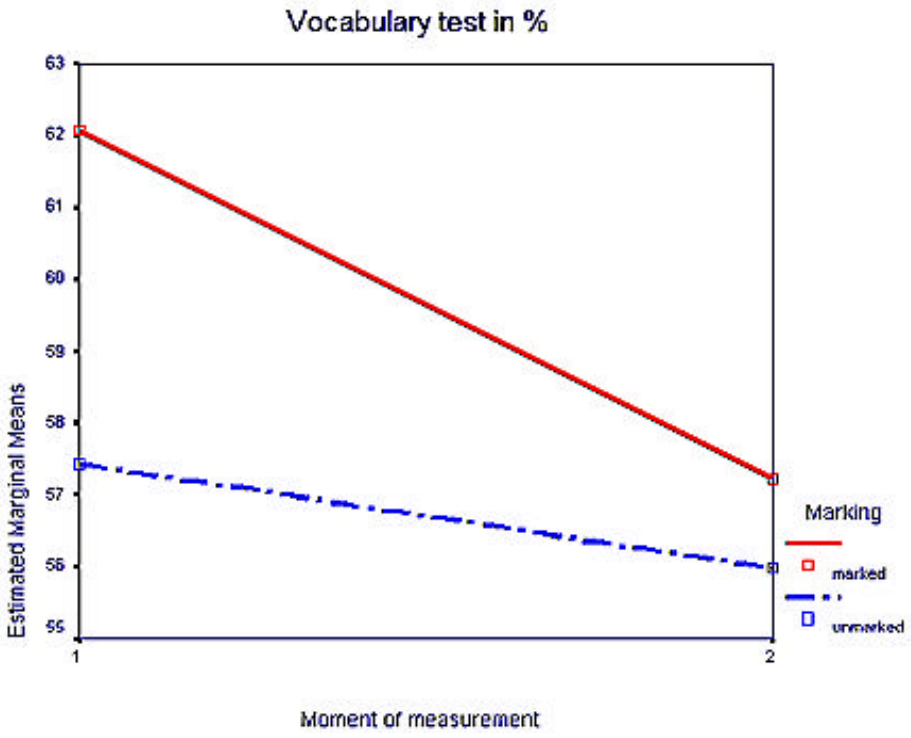

Figure 5. Vocabulary Test in Percentage, 2 Within-Subjects Variables

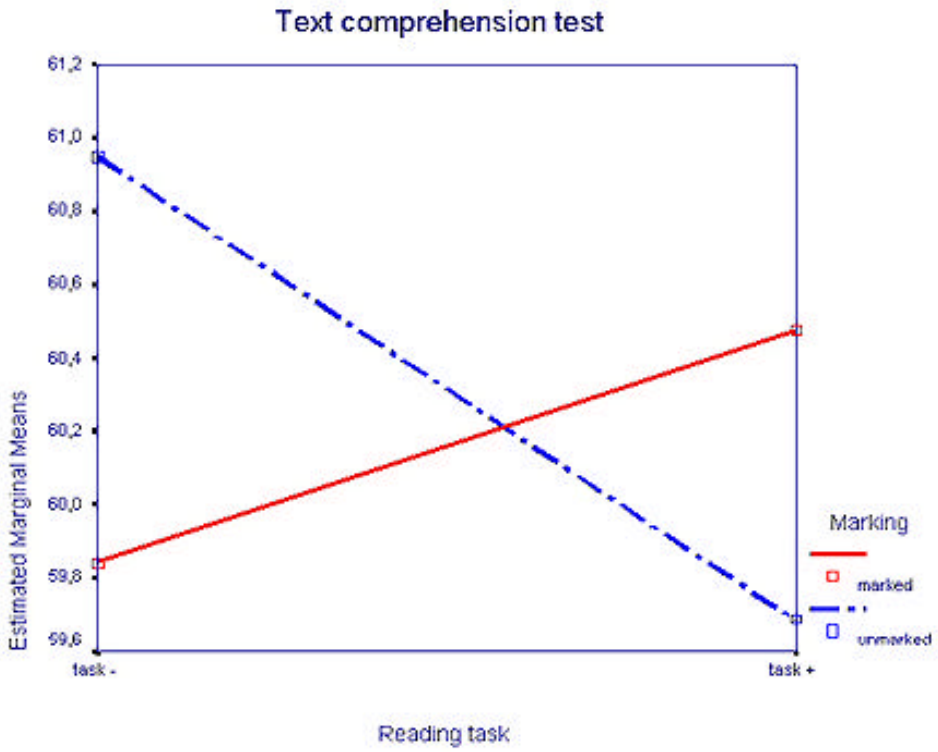

Figure 6. Text Comprehension Test in Percentage, Within-Subjects and Between Subjects Variable

NOTES

1. Other factors seem to have a positive influence on incidental vocabulary learning. Hulstijn et al. (1996) give a detailed overview of these factors with full bibliographical references.
2. None of the log-times of glosses proved to be enormously long. It is possible that the students left the pop-up window open while rereading the sentence or while copying vocabulary on paper. However, these activities can be categorised under "consultation" of a gloss.
3. In text 1 about the diamond industry: *cautioler* (to guarantee), *être hardelé à* (to be authorised to), *le desti* (the challenge), *pourver* (to push to extremes). In text 2 about the human capital in business: *éloxer* (to remove), *mettre en devore* (to incite), *rummérer* (to rime), *berluter* (to encounter).
4. The vocabulary tests have a very high reliability index (Cronbach's alpha):
   Text 1, immediately after reading, α = . 8476; Text 1, delayed, α = .8677, *N* of items = 38
   Text 2, immediately after reading, α = . 8208; Text 2, delayed, α = .8015, *N* of cases = 60
   When applying the filter of pre-knowledge:
   Text 1, immediately after reading, α = . 8423; Text 1, delayed, α = .8606,
   N of items = 35, *N* of cases = 60
   Text 2, immediately after reading, α = . 8254; Text 2, delayed, α = .8133,
   N of items = 30; *N* of cases = 60
5. The Cronbach's alpha of the free recalls:
   Text 1, α = .5709; *N* of items = 22
   Text 2, α = .5862; *N* of cases = 30

6. These authors propose to confront the students with an auditory signal during writing, on which they are asked to produce an oral reaction (i.e., say "stop").
7. The results of these general questions and interviews have been analysed but did not reveal any relevant differences. I have therefore chosen not to report them within the scope of this article.
8. The fact that the free recall was performed before the students took the comprehension test did not have a facilitating effect on the latter, since correlations between both tests are not significant (Pearson product moment coefficients: marked, .171; unmarked, .144).
9. All the results were analysed with a General Linear Model (GLM), Repeated Measures (unless otherwise indicated). This is a Repeated Measures Analysis of Variance (an ANOVA for Repeated Measures). Using the GLM-procedure, one can test null hypotheses about the effects of both the within-subjects and the between-subjects factors.
10. These results are confirmed when taking into account the percentage of words clicked. The results show that marking and reading task have a significant effect on the student's clicking behaviour: marking: F (1,58)=13.992, $p < .05$; reading task, F (1,58)=14.40, $p < .05$). There is no interaction effect between the within-subjects variable and the between-subjects variable. F (1,58)=1.576, $p < .05$.
11. For all of the students involved, the vocabulary test of the first session (marked or unmarked) was completely unexpected. It is possible that in the second session, the students might have expected a similar test. A GLM Repeated Measures Model including "session" and "time" as within variables and "reading task" as a between variable yields no significant effect of session, F (1,58)=1.53, $p = .221 > .05$; no interaction effects were established, only an effect of time was found, F (1,58)=8.71, $p = .005^* < .05$.
12. When only taking into account the non-existent words, this results in a (fairly reliable) test where pre-knowledge is absolutely excluded. Reliability:
    Text 1, Cronbach's $\alpha = .7426$, N of items = 4
    Text 2, Cronbach's $\alpha = .5773$, N of cases = 60
    The results on this test have been analysed with a GLM repeated measures and confirm the results of the overall vocabulary test. The results show that on the one hand, marking has no significant effect on vocabulary learning, F (1,58)= .311, $p > .05$. On the other hand, time and reading task do have a significant effect on vocabulary learning: time, F (1,58)=17.75, $p < .05$; reading task, F (1,58)=26.40, $p < .05$. There are no interaction effects: marking*reading task, F (1,58)=2.10, $p > .05$; marking*time, F (1,58)=1.58, $p > .05$; time*reading task, F (1,58)=.219, $p > .05$; marking*task*time, F (1,58)=2.62, $p > .05$).
13. When computing a Pearson product moment coefficient, the correlation between text comprehension and vocabulary learning obtained is not all that strong: For the marked condition a correlation of .172 (not significant) was obtained and for the unmarked condition .333** (= significant at the .05 level).
14. Thirty seven out of 60 students took notes: 12 in the group with the general reading task and 25 in the group with the specific reading task.

ACKNOWLEDGEMENTS

I am very grateful to Gert Rijlaarsdam and Luuk Van Waes for reading and commenting on various drafts of this paper. I am also indebted to Mariëlle Leyten for her help in analysing the recalls. Finally, I would also like to thank the anonymous reviewers for their valuable comments and suggestions.

ABOUT THE AUTHOR

Isabelle De Ridder graduated in 1993 in romance languages and has been a foreign language teacher ever since. She is currently a fourth year PhD student at the University of Antwerp (UFSIA). Her research focuses on online reading in a foreign language and screen design implications.

E-mail: isabelle.deridder@ua.ac.be

REFERENCES

Benítez, L., Castrillo, J. M., Cerezal, N., & Suárez, C. (1988). *Reading tasks.* Harlow, UK: Longman.

Black, A., Wright, P., Black, D., & Norman, K. (1992). Consulting on-line dictionary information while reading. *Hypermedia 4*(3), 145-169.

Brett, P. (1997). A comparative study of the effects of the use of multimedia on listening comprehension. *System 25*(1), 39-53.

Brett, P. (1998). Using multimedia: A descriptive investigation of incidental language learning. *Computer Assisted Language Learning 11*(2), 179-200.

Chun, D. M., & Plass, J. L. (1996). Effects of multimedia annotations on vocabulary acquisition. *Modern Language Journal 80*(2), 183-198.

Chun, D. M., & Plass, J .L. (1997). Research on text comprehension in multimedia environments. *Language Learning & Technology 1*(1), 60-81.

Craik, F., & Lockhart, R. S. (1972). Levels of processing: A framework for memory record. *Journal of verbal learning and verbal behaviour 11,* 67-84.

De Ridder, I. (1999). Are we still reading or just following links? In K. Cameron (Ed.), *CALL and the learning community* (pp. 195-116). Exeter, UK: ELM Bank Publications.

De Ridder, I. (2000). Are we conditioned to follow links? Highlights in CALL-materials and their impact on the reading process. *Computer Assisted Language Learning 13*(2), 183-195.

De Ridder, I., & Van Waes, L. (2000). Gemarkeerde hyperlinks en hun invloed op het leesproces, de woordenschatverwerving en het tekstbegrip [Marked hyperlinks and their influence on reading process, vocabulary acquisition and text comprehension]. In N. Ummelen, R. Neutelings, & F. Maes (Eds.), *Over de grenzen van de taalbeheersing. Lezingen gehouden op het VIOT congres in Delft* [Over the borders of language competence. Proceedings of the VIOT conference in Delft, The Netherlands] (pp. 381-391). Den Haag, The Netherlands: SDU.

Dillon, A. (1992). Reading from paper versus reading from screens: A critical review of the empirical literature. *Ergonomics 35,* 1297-1326.

Drum, P. A., & Konopak, B. C. (1987). Learning word meanings from written context. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 73-87). Hillsdale, NJ: Erlbaum.

Dyson, M. C., & Kipping, G. J. (1998). The effects of line length and method of movement on patterns of reading from screen. *Visible Language 32*(2), 151-181.

Ellis, R., Tanaka, Y., & Yamazaki, A. (1994). Classroom interaction, comprehension, and the acquisition of word meanings. *Language Learning 44*(3), 449-491.

Fisher, D. L., & Tan, K. C. (1989). Visual displays: the highlighting paradox. *Human Factors 31*(1), 17-30.

Freebody, P., & Anderson R. C. (1983a). Effects of vocabulary difficulty, text cohesion, and schema availability on the reading comprehension. *Reading Research Quarterly 18,* 277-305.

Freebody, P., & Anderson R. C. (1983b). Effects of text comprehension of differing proportions and locations of difficult vocabulary. *Journal of Reading Behaviour 15,* 19-40.

Frenckner, K. (1990). Legibility of continuous text on computer screens -- a guide to the literature (TRITA-NA-P9010). Stockholm: Royal Institute of Technology.

Fyfe, R., & Mitchell, E. (1985). *Reading strategies and their assessment.* Oxford, UK: The NFER-NELSON Publishing Company.

Godwin-Jones, R. (2000, June). Better Web-development: designing for user customization. Paper presented at the CALICO conference, Tucson, Arizona.

Hayes, J. R., & Hatch, J. A. (1999). Issues in measuring reliability. *Written Communication 16*(3), 354-367.

Holley, F. (1973). A study of vocabulary learning in context: The effect of new-word density in German reading material. *Foreign Language Annals 6,* 339-347.

Huckin, T., & Coady, J. (1999). Incidental vocabulary acquisition in a second language. *Studies in Second Language Acquisition 21*(1), 181-193.

Hulstijn, J. H. (1992). Retention of inferred and given meanings: Experiments in incidental vocabulary learning. In P. J. L. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 113-125). London: Macmillan.

Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use and reoccurence of unknown words. *The Modern Language Journal 80*(3), 327-339.

Jacobs, G. M., Dufon, P., & Fong, C. H. (1994). L1 and L2 vocabulary glosses in L2 reading passages: Their effectiveness for increasing comprehension and vocabulary knowledge. *Journal of Research in Reading 17,* 19-28.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice Hall.

Kellog, R. T., & Mueller S. (1993). Performance amplification and process restructuring in computer-based writing. *International Journal of Man-Machine Studies 39*(1), 33-50.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* Cambridge, UK: University Press.

Knight, S. (1994). Dictionary: The tool of last resort in foreign language reading? A new perspective. *The Modern Language Journal 78,* 285-299.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: additional evidence for the input hypothesis. *Modern Language Journal 73*(4), 440-464.

Liu, M., & Reed, W. M. (1995). The effect of hypermedia-assisted instruction on second-language learning. *Journal of Educational Computing Research 12*(2), 159-175.

Lomicka, L. (1998). "To gloss or not to gloss": An investigation of reading comprehension online. *Language Learning and Technology 1*(2), 41-50.

Muter, P. (1996). Interface design and optimization of reading of continuous text. In H. van Oostendorp & Sj. de Mul (Eds.). *Cognitive aspects of electronic text processing* (pp. 161-180). Norwood, NJ: Ablex Publishing Corporation.

Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly 20,* 233-253.

Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York: Newbury House.

Nielsen, J. (2000). *Designing Web usability*. Indianapolis, IN: New Riders Publishing.

Nuttall, C. (1987). *Teaching reading skills in a foreign language*. London: Heinemann Educational Books.

Roby, W. B. (1999). What's in a gloss? *Language Learning and Technology 2*(2), 94-101. Retrieved August 13, 2001, from http://llt.msu.edu/vol2num2/roby/

Schmalhofer, F., & Glavanov, D. (1986). Three components of understanding a programmer's manual: verbatim, propostional, and situational representations. *Journal of Memory and Language 25*, 279-294.

Scott Grabinger, R., & Osman-Jouchoux, R. (1996). Designing screens for learning. In H. van Oostendorp & Sj. de Mul (Eds.), *Cognitive aspects of electronic text processing* (pp. 181-212). Norwood, NJ: Ablex Publishing Corporation.

Stark, H. A. (1990). What do readers do to pop-ups, and pop-ups do to readers? In R. McAleese & C. Green (Eds.), *Hypertext: State of the art* (pp. 2-9). Oxford, UK: Intellect Limited.

Sternberg, R. J., & Powell, J. S. (1983). Comprehending verbal comprehension. *American Psychologist 38*, 878-893.

Sternberg, R. J. (1987). Most vocabulary is learned from context. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 89-105). Hillsdale, NJ: Erlbaum.

Vermunt, J. D., & Van Rijswijck, F. A. (1987). *Inventaris leerstijlen voor het Hoger Onderwijs* [Inventory of learning styles for higher education]. Tilburg, The Netherlands: Katholieke Universiteit Brabant.

Wesdorp, H. (1981). *Evaluatietechnieken voor het moedertaalonderwijs* [Evaluation techniques for mother tongue education]. 's Gravenhage, The Netherlands: Staatsuitgeverij.

West, M. (1941). *Learning to read a foreign language*. Harlow, UK: Longman.

Widdowson, H. G. (1978). *Teaching language as communication*. Oxford, UK: Oxford University Press.